A Study on Characterization of Gate Oxide Shorts Using Non-split Model

Chua Yong Moh, Abu Khari bin A'ain
Fakulti Kejuruteraan Elektrik,
Universiti Teknologi Malaysia,
81310 Skudai, Johor, Malaysia
e-mail: yanmao@hotmail.com, yanmao@time.net.my

Abstract- The integrity of Gate Oxide Shorts (GOS) model is a key factor as quality and reliability indicator of CMOS. Gate Oxide defects in MOS transistors can be considered as the layout and technology dependent failures for which logic fault models are not always available, requiring electrical models to simulate the defect characteristics. Previously the GOS have been modeled with split transistors technique using two minor transistors and lumped elements. However, it is problematic to study minimum size transistors affected by GOS failures using the existing unidirectional split model as the channel length is designed at minimum size in particular technology process. This paper presents a study to compare and correlate between split model and non-split model of GOS.

Keywords
Gate Oxide Shorts (GOS), unidirectional model, non-split model, technology process, CMOS

I. INTRODUCTION

Gate oxide shorts in MOS transistors are the result of manufacturing process or material defects. They usually have limited extent in comparison with the dimensions of MOSFET and randomly distributed over the silicon wafers [1]. These shorts exhibit diverse electrical properties and therefore can be responsible for IC faults. Break down of the gate oxide layer result of lithography defects on mask, deviation on thin oxide growing and field failures due to voltage overstress can be considered as the main reasons for gate oxide shorts.

Lithography defects such as airborne particles or mask damages may be responsible for lack of gate oxide layer under small portion of polysilicon gate. On the other hand, the electrical breakdown of the gate oxide layer may be induced by a high electric field that initiates injection of electrons from the electrode into the gate oxide and by the subsequent impact ionization. Thus missing spot occurs over the MOS transistor channel. A missing gate oxide spot will be filled by gate material, and that the dopant atoms from the gate will be able to penetrate to the substrate of the device.

The lithography defects introduce pinholes in gate oxide which contribute to nearly zero breakdown fields to conduct a significant gate current in the defective MOS transistor [1]. Figure 1 shows an example of GOS defect in a MOSFET transistor. As for electrical breakdown defects, the SiO_2 layer that contain defects are characterized by smaller breakdown field. It introduces weak spots in the oxide, and the energy dissipated during breakdown of gate oxide layer may produce significant changes in the MOS structure. Because of the small sizes of the oxide defects, the Si filaments that short the polysilicon gate to the Si surface will have resistance. However, the value of the resistance varies depends on many factors such as defect size, oxide thickness, and doping densities in the polysilicon gate and diffusion regions.

The gate oxide shorts in MOS transistor may be located in various regions of the gate oxide layer. When a gate oxide short occurs in the overlap region, the gate becomes shorted either to the drain or to the source region of the transistor. Whereas when the gate oxide short occurs within the MOSFET channel and far from the overlap regions, it is being recognized as gate to channel short.

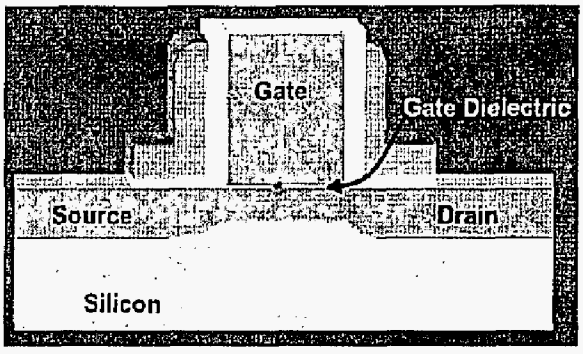

Figure 1. Gate Oxide Short

This paper is organized in the following way. Section 2 describes the Gate Oxide Shorts models. Section 3 presents the simulation results and discussion. Section 4 concluded the paper.

II. GATE OXIDE SHORTS MODELS

A fault is a deterministic, discrete change in circuit behavior whereas defect is non-deterministic and random in nature [2]. Depending on the defect location, the GOS can be categorized into two types. The first type of defect which connects the gate to drain/source diffusion region is referred as an ohmic defect and exhibits a linear behavior in the IG versus VGS characteristics. The second type of defect which connects the gate to channel is referred as non-ohmic defect and exhibits a non-linear behavior in the IG versus VGS characteristics [3].

A simple linear model has been proposed since year 1980's and it is based on the addition of short resistance between the gate and drain/source electrodes [1], [4]. The value of the resistance depends on many factors such as defect size, oxide thickness, and doping densities in the polysilicon gate and diffusion region. Due to special feature in CMOS processes, GOS in n-transistor has different properties from those in p-transistor and thus should be modeled differently. The difference between a GOS in a n-transistor and in a p-transistor is that the GOS in p-transistor has a pn-junction diode in series with short resistance [5], [6], [7]. The main advantage of this model is that the length of the transistor is preserved and implying that the model can handle transistors which are designed in minimum size in the particular process technology. However, the model is not able to represent the non-linear behavior which is presented due to the shorts between the gate and the channel. Thus, through out the years, researchers have been introducing various types of fault models to model the gate to channel shorts.

A. Bi-Directional Model

In order to analyze the gate to channel defect through electrical SPICE simulation, a fault free transistor had been modeled based on lumped-elements. In the bi-directional model, the non-defective channel is represented by a bi-directional array of MOS transistors with m lines and n columns which consist of $n+1$ transistors and n nodes per line and m-1 transistors and m nodes per column [1], [6]. This is shown in Figure 2.

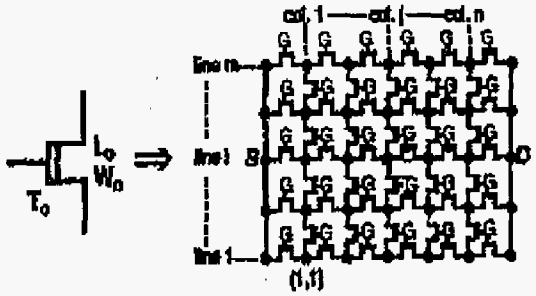

Figure 2. Lumped Element Bi-directional Model

In this model, the pinhole is modeled by a resistor, Rgos between the common gate and one of the internal nodes of the network. The defect resistance and location can then be varied through the value of Rgos and position in the network. There are three major parameters in this model which are the location of the faults, the size of the faults and resistance of the Rgos.

B. Unidirectional Model

From bidirectional model, researchers observed that the Vin and Vout characteristics are not really sensitive to the transversal location of the short [6]. Hence, unidirectional model has been introduced in order to reduce the complexity of the simulations.

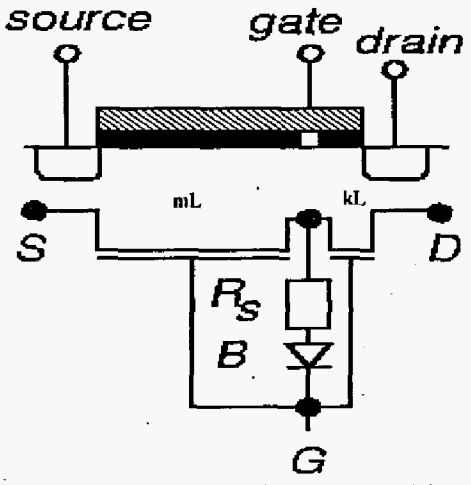

Figure 3. Unidirectional Model

In an unidirectional fault model, a defective MOS transistor is modeled by 3 components, a rectifying barrier and two minor transistors in which the faulty transistor can be seen as a device split into two minor transistors [8], [9]. The junction between the gate and channel is modeled at circuit level by a rectifying barrier between the gate and the channel. The barrier is characterized by a threshold voltage, a breakdown voltage, forward resistances and backward resistances [8].

The contact resistance of the short is modeled by Rgos. The point where the short is located splits the transistor in two sections, behaving as two smaller transistors as shown in figure 3.

A short location can be determined by its distance kL from drain and its distance of mL from source where L is the length of the fault free transistor. In this case, $k+m=1$ and a rectifying barrier exist between the substrate and the gate. This is shown in Figure 3. Let β be the transconductance factor of the fault free transistor and β_1 and β_2 correspond to the two smaller transistors The overall transconductance of the faulty transistor is same as the fault free transistor[8], [9].

$$\beta_1 = \beta/m$$

$$\beta_2 = \beta/k$$

$$\beta_1 \beta_2/(\beta_1+\beta_2) = \beta$$

C. Non Linear Non Split Model

As the transistor size is continually being scaled down and most of the digital ICs are designed in minimum size, it is problematic to study minimum size transistors affected by GOS failures using the existing unidirectional split model. This is due to the fact that the channel length is designed at minimum size in the particular technology process. Thus, non-linear non-split model has been proposed [3]. The objective of the non split model is to preserve the length of the original transistor and to introduce non-linearity in the model in order to properly represent the behavior of the GOS. Figure 4 shows an example of the new non-split model.

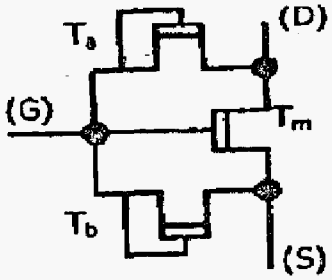

Figure 4. Non-split Model

The non split model is derived using flow characteristics approach by adding non-linear components to the original transistor. In this case, two transistors are added to the original transistor which each of them is responsible for the negative drain current, $-I_d$ and leakage current, I_g of the faulty transistor. The model is constructed by first modifying the width of the original transistor; followed by inserting two additional transistors between the gate and drain/source of this original transistor. The model parameters (W_m, W_a/L_a, W_b/L_b) are then tuned to reproduce GOS defects of various resistance, location and size [3].

III. SIMULATION RESULTS AND DISCUSSION

In this simulation, transistors of 1.2 micron technology were used as to verify the correlation between both unidirectional model and new non-split model. The reasons of using 1.2 micron technology are because of the well established process technology and capability of the existing unidirectional split model in modeling the GOS defects behavior. The tuning process can be done based on the existing unidirectional split model as long as the transistors are not in their minimum sizes. The new non-split model simulation results can be analyzed separately for each transistor. First of all, the first transistor which mainly contributes negative current to the transistor when |Vgs| > |Vds| is observed. Transistor 2 is in-charge of the drain current in the faulty transistor. The gate leakage current is mainly controlled by transistor 3. Figure 5 a) and 5 b) show the current in both the split model and non-split model respectively.

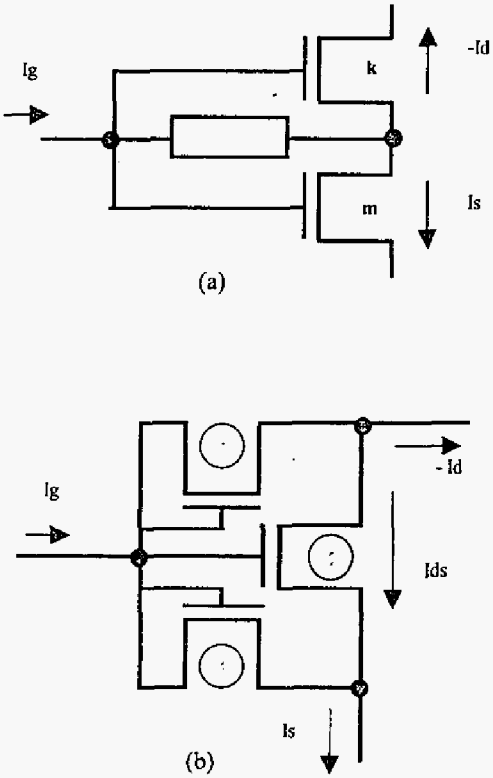

Figure 5. (a) Current Flow in Split Model (b) Current Flow in Non-split Model

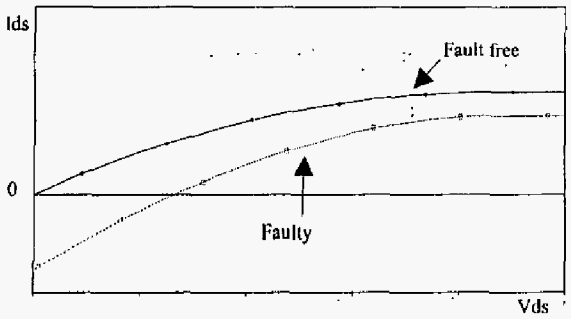

Figure 6. Ids vs Vds characteristic of a faulty and fault free nmos

Generally, a transistor with GOS defect will have smaller Ids value than fault free transistor for a given Vgs. Furthermore, the existence of the negative drain current will occur for smaller Vds value. The comparison between fault free and faulty transistor characteristic is shown in figure 6.

From the observation of figure 7, for fault which is closer to drain, the –Id will be greater compared to the fault located near to the source. This is because the current of transistor k as shown in figure 5a) will be in the saturated region. In addition, the internal resistance for the current path is lesser when fault is located near to the drain. This can be understood by observing the changes of the W/L ratio of the unidirectional split model. When the fault is near to the drain, the W/L of the transistor k (which is connected as drain in the split model) is getting larger. Thus, the resistance of this current path is getting smaller as the fault move towards the drain. This phenomenon can also be described using the new non-split model. In the new model the leakage current, Ig is the combination of the Id_1 and the Id_3 (which is mainly contributed for faulty gate current).

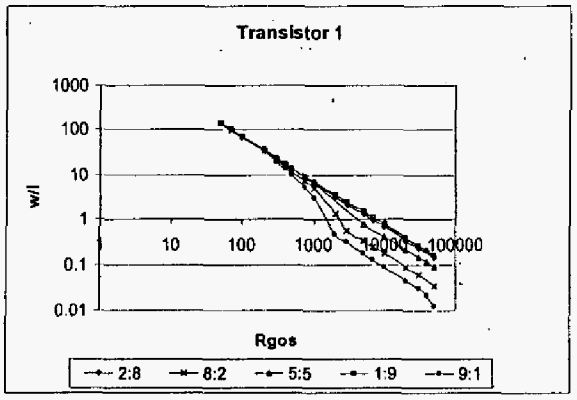

Figure 7. W/L ratio of transistor 1 in various Rgos and location for non-split model

Figure 7, 8 and 9 show simulation results which correlate the characteristic between split model and non-split model. The ratio 2:8 for example means k:m for the defect location which refers to the split model. From the graph in figure 7, notice that for large Rgos value, fault which is nearer to drain the W/L ratio of transistor 1 is much greater than fault which is near to the source. This shows that fault; which is near to the drain will drive more negative current than fault, which is near to the source. When the shunt resistance reduces, the effect of defect location is very small. When |Vgs|> |Vds|, the high leakage current due to low Rgos does guarantees high – Id no matter where the fault location is. Thus, the W/L ratio is almost the same when the Rgos is small regardless the changes of fault location.

Transistor 2 on the other hand, plays an important role in Ids current flow in the circuit. It controls the current flow by adjusting the W/L ratio of transistor 2. Figure 8 shows that fault which is near to the source will have larger W/L ratio as the Rgos reduces.

It is obvious that, the responses saturated at both ends. For Rgos larger that 50 kΩ, the fault locations make no difference to the W/L ratio. It is because the high resistance approximates an open circuit to the gate oxide shorts. On the other hand, low resistances guarantee the existence of the faults in the circuit.

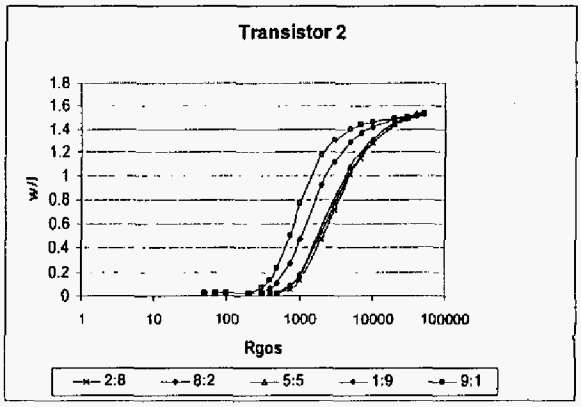

Figure 8. W/L ratio of transistor 2 in various Rgos and location

In CMOS, Ids of the nmos allow the discharge of the voltage, thus greater w/L ratio in transistor 2 allows larger Ids to flow through the faulty transistor. In a particular shunt resistance Rgos, fault which is located near to the source will have a larger W/L ratio compared to fault located near to the drain. It is because; the V_{gs} will be sunk by the R_{gos} and the internal resistance of the faulty transistor. When $V_{gs} \geq$ Vtho, the transistor will operates in the saturated region as Vds will be greater than Vgs-Vtho. The low W/L ratio will greatly reduce the faulty transistor ability to discharge. Thus, it is necessary to find out the faulty region where the fault can be

detected using logical test. Thus, the point where the graph saturated (minimum value) at can be referred as the range of 'hard' faults in terms of Rgos.

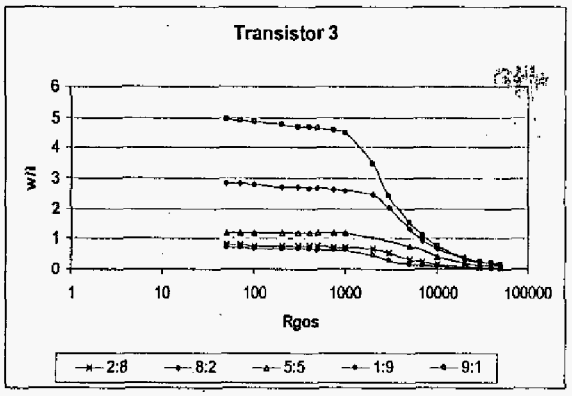

Figure 9. w/L ratio of transistor 3 in various Rgos and location

From the graph in Figure 9, it can be easily seen that for fault which is closer to the drain, it tend to saturate near the value 1. For transistor 3, the larger the w/L ratio will allows larger gate current in the circuit. For fault which is near to the source, the W/L ratio is larger. It is due to the low resistance in the current path when the fault is near to the source. When the Rgos is 1 kΩ, the graph tends to reach their saturate values. And thus, it is seen that the fault tends to cause the circuit to logically fail when Rgos reaches 1 kΩ. It is thus, say that the hard fault boundary will be at Rgos = 1 kΩ. The non-linear increment of the W/L ratio for different fault locations seems to be related to the non-linear part in the graph of the transistor 2. Both of them imply the non-linear effects of the GOS defects.

IV. CONCLUSION

In this paper, a new non-linear non-split model is used for gate to channel defects. The objective is to emulate the original length of the transistor in order to handle minimal length transistors. Based on the analysis of the correlation between unidirectional split model and the non-split model, we can conclude that the changing of a parameter in the split model would have correspondent changes in the new model and it follows certain trend as parameter shifted. For a given Rgos boundary, defects near to the drain may have functional disability. Whereas, defects near to the source may introduce performance degradation as its leakage current increases the power dissipation of the faulty transistor.

Future work will be concentrated on the simulations of GOS in various process technologies in order to obtain the trend of the defect from one process technology to another. It is also possible to come out with a conversion table between the parameters of the existing unidirectional model and the new non-split model.

V. REFERENCES

[1] Marek Syrzycki, "Modeling of Gate Oxide Shorts in MOS Transistors", IEEE Transactions on Computer-Aided Design, VOL. 8, pp. 193-202, March 1989.

[2] Robert C. Aitken, "Nanometer Technology Effects on Fault Models for IC Testing", IEEE, pp.46-51, Nov 1999.

[3] M. Renovell, J.M. Galliere, F. Azais, Y. Bertrand, "Modeling Gate Oxide Short Defects in CMOS Minimum Transistors", Proceedings of the Seventh IEEE European Test Workshop, 2002.

[4] Hong Hao, Edward J. McCluskey, "On The Modeling And Testing Of Gate Oxide Shorts In CMOS Logic Gates", Int. Workshop on Defect and Fault Tolerance on VLSI System, pp. 161-174, 1991.

[5] J.Segura, C. De Benito, A. Rubio, C.F. Hawkins, "A Detailed Analysis of GOS Defects in MOS Transistors: Testing Implications at Circuit Level", Int. Test Conf., pp. 544-551, 1995.

[6] M. Renovell, J.M. Galliere, F. Azais, Y. Betrand,, "Boolean and Current Detection of MOS Transistor with Gate Oxide Short", Int. Test Conf., pp. 1039-1048, 2001.

[7] M. Shaheen Sayeed, Samiha Mourad, "Gate To Channel Shorts In PMOS Devices: Effects On Logic Gate Failures", IEEE, pp.211-214, 1998.

[8] J. Segura, A. Rubio, J. Figueras, "Analysis And Modeling of MOS Devices With Gate Oxide Short Failures", Int. Symp. Circuits and System, pp. 2164-2167, 1991.

[9] R. Rodriguez-Montanes, J. Segura, V. Champac, J. Figueras, J. Rubio, "Current vs Logic Testing of Gate Oxide Short, Floating Gate and Bridging Failures", Int. Test Conf., pp. 510-519, 1991.

[10] Abdullah Yassine, Karsten Wieczorek, Kola Olasupo & Volker Heinig, "A Novel Electrical Test to Differentiate Gate-to-Source/Drain Silicide Short from Gate Oxide Short", IRW Final Report, 2000.